

ABSTRACT

Performance parameters are obtained for distributed processing units. The performance parameters include a utilization value, of each distributed processing unit. Respective weights are obtained for the distributed processing units by applying a mapping function to the utilization values. The respective weights are used for weighted round-robin load balancing of work requests upon the distributed processing units. In one implementation, a set of utilization values of the processing units are collected in response to a periodic heartbeat signal, a set of weights are produced from the set of utilization values, a distribution list is produced from the set of weights, and the distribution list is randomized repetitively for re-use during the heartbeat interval.